Lingua 122 (2012) 145–163

Contents lists available at ScienceDirect

Lingua

journal homepage: www.elsevier.com/locate/lingua

A phonetically based metric of sound similarity☆

Jeff Mielke *

Department of Linguistics, University of Ottawa, Arts 401, 70 Laurier Ave East, Ottawa, ON K1N 6N5, Canada

ARTICLE INFO

Article history:
Received 21 January 2011
Received in revised form 20 April 2011
Accepted 21 April 2011
Available online 24 June 2011

Keywords:
Phonetic similarity
Phonological similarity
Ultrasound
Electroglottographym
Airflow
Acoustic

ABSTRACT

This paper examines similarity measures based on acoustic and articulatory data from a set of crosslinguistically frequent consonants and vowels, and compares this phonetic similarity with measures of phonological similarity that are based on the crosslinguistic patterning of phonemes associated with these sounds.

1. Introduction

A basic question in phonological theory is *Why is the class of obstruents active in many different languages?* Many similarly phrased questions can be asked about other groups of sounds. The leading answer to this question may once have been *Because the feature [sonorant] is in Universal Grammar*. In the absence of this UG-based explanation, a comparable answer would be *Because obstruents are phonetically similar to each other*. Supporting this second answer requires a means of quantifying phonetic similarity. Phonetic similarity is frequently invoked for explaining a wide range of phonological observations, but to this day no objective metric of phonetic similarity is widely available.

The purpose of the project described in this paper is to provide a resource for quantifying phonetic similarity, distinguishing articulatory, acoustic, and perceptual similarity from each other and from phonological notions of similarity, such as those based on features (e.g. Frisch, 1996; Frisch et al., 2004; Kondrak, 2003, *et seq.*) or on phonological patterning. This paper examines similarity measures based on acoustic and articulatory data from a set of crosslinguistically frequent consonants and vowels, and compares this phonetic similarity with measures of phonological similarity that are based on the crosslinguistic patterning of phonemes associated with these sounds.

☆ Collection of the phonetic data reported here was made possible by Diana Archangeli and Natasha Warner and supported by James S. McDonnell Foundation grant #220020045 BBMB to Diana Archangeli and funding from the College of Social and Behavioral Sciences at the University of Arizona. Data analysis was supported by SSHRC grant #410-2007-0735, "Measuring the Phonetic Similarity of Speech Sounds" and funding from the University of Ottawa Faculty of Arts and the University Research Committee. The collection of P-base was supported by a Presidential Fellowship from the Ohio State University. Research assistants for data analysis at the University of Ottawa were Tanya Awad, Daryl Chow, Chris Delaney, Tabish Ismail, Lyra Magloughlin, Kayleigh Markell, Joseph Roy, Seiko Sagae, Ricardo Tabone, Agnesa Terziyska, and Marie-Claude Tremblay. This project has benefited from comments from Natasha Warner, Osamu Fujimura, Shigeto Kawahara, Søren Wichmann, an anonymous reviewer, and audience members at NELS 40. All mistakes are mine.

* Tel.: +1 6135625800x2701.
E-mail address: jmielke@uottawa.ca.

0024-3841/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.lingua.2011.04.006

The original motivation for creating the similarity metric was to investigate the role of phonetic similarity in determining the sets of segments that are involved in sound patterns. It has long been known that the phonologically active classes appearing in languages' sound systems are not randomly assembled, but reflect general crosslinguistic tendencies. Distinctive feature theory (Chomsky and Halle, 1968; Clements, 1985; Clements and Hume, 1995, *inter alia*) was put forth as a model of the phonetic parameters used to form contrasts and define natural classes. More recently, it has been argued that an all-purpose feature model is in conflict with many attested phonologically active classes, including recurrent ones (Mielke, 2008a), and that the preference for certain classes is driven by physiological and perceptual factors and their role in diachronic change (Ohala, 1981; Blevins, 2004, *inter alia*). The goal of this project is to provide tools for investigating the role of ubiquitous phonetic factors in accounting for which sounds tend to pattern together and which sounds tend not to.

The purpose of this project is not to replace an all-purpose phonological model with an all-purpose phonetic model. However, the study of phonology does deal with a wide range of observations that by their nature are large-scale and language-independent, and make use of broad "all-purpose" statements, such as markedness generalizations, general phonetic explanations for typological patterns, and the types of phenomena that are often accounted for in terms of distinctive features and other primitives of phonological theory. These observations have often received language-independent phonetic and phonological accounts. One of the reasons for focusing on phonetically based explanations for phonological patterns (in spite of the fact sound patterns are also influenced by social, cognitive, lexical, and grammatical factors), is that statements about phonetic tendencies apply generally to all spoken languages, and the other factors vary too much for general statements to be made as easily. This idea of phonetics as a common denominator is the basis for historical phonetic accounts of phonological patterns (Ohala, 1981, 1983, 1993, 1994; Blevins, 2004). Formal primitives such as distinctive features have been useful in investigating crosslinguistic patterns, because they are designed to apply to all languages. This project aims to provide a measure of similarity that does not rely on formal primitives. Nevertheless, there is no substitute for accounting for phonological patterns with phonetic data from the languages where the patterns are observed, and to use particular phonetic data produced by speakers of another language as a stand-in is obviously still a simplification.[1]

This project aims to provide a more direct route from phonetic observations to phonological patterns, unmediated by formal primitives, but it still does not avoid idealization. All of the sounds described here were produced in intervocalic position in a laboratory by linguists whose native language is English. There are many important ways in which a particular consonant or vowel produced under these circumstances does not adequately represent a consonant or vowel in a particular language that could potentially be identified with it. Any segmental properties that are specific to other languages, prosodic positions, vowel contexts, speech styles, or anything like that would not be represented in the data resulting from this project.[2] However, many phonologically relevant properties are expected to hold up. For example, the [t]s produced here are all produced as voiceless coronal stops, as are most of the [t]s of the world's languages. So while not every phoneme described as /t/ is pronounced the same way, it is still useful to have some phonetic data points for segments described as [t], as compared to other segments, as an alternative to using a bundle of distinctive features or a three-part articulatory description as the common denominator. For many of the phonetic properties linguists are interested in, and that are relevant for language-independent statements, phonetic measurements of the articulatory and acoustic properties of speech sounds produced under specific circumstances bear non-trivial relevance.

The presentation of the data in this paper involves a specific set of measurements of phonetic data that could be measured in numerous different ways. As the measurements are discussed in the next section, advantages and limitations and other alternatives will be described where relevant.

In addition to the phonetic similarity data, this paper describes a metric of phonological similarity based on shared phonological patterning in a database of sound patterns. This database is organized in terms of phonemes. Using phonetic data of the kind reported here to address data involving phonemes requires a simplifying assumption, specifically the association of a phoneme with a particular allophone of that phoneme. This is similar to the assumptions that linguists use when making use of resources such as segment inventory databases like UPSID (Maddieson and Precoda, 1990), which uses IPA symbols to describe phoneme inventories in a range of languages that do not all realize these phonemes in the same ways. See Simpson (1999:349) for a critique the use of UPSID in these terms and see also Mielke (2008b) for a discussion of these issues. For the same reasons, the phonetic representations of the type described here clearly do not interface directly with language-specific phonological representations, which involve various phonemic oppositions and allophonic patterns. What allows this approach to make sense is the assumption that a phoneme that is described with a particular IPA symbol will have some connection with a speech sound produced by a linguist reading the same symbol. There is obviously not a direct connection between a phonetic representation resulting from an English-speaking linguist in a laboratory and a phonology-phonetics mapping in a different language, but the potential connections are also obvious, and no more remote than the connection between a set of distinctive features and a phonological representation in a particular language (until shown otherwise).

[1] Note that using phonetic data from the same language also involves a simplification in may cases: if the role of phonetic effects is primarily in diachrony, then the most appropriate phonetics experiment would require access to speakers of an earlier stage of the same language.

[2] See, e.g. Sproat and Fujimura (1993) on the importance of prosodic position for segment realization.